

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE May 22, 2001		3. REPORT TYPE AND DATES COVERED Final Report - 9/28/98 - 9/27/00	
4. TITLE AND SUBTITLE Genetic Engineering of Artificial Proteins Containing Non-Natural Amino Acids				5. FUNDING NUMBERS G - DAAG55-98-1-0518	
6. AUTHOR(S) David A. Tirrell					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) California Institute of Technology Dept. of Chemistry & Chemical Engineering - 210-41 Pasadena, CA 91125				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211				10. SPONSORING / MONITORING AGENCY REPORT NUMBER 39211.1-LS	
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
12 a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				12 b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This project has led to new approaches to protein engineering through incorporation of non-natural amino acids. Protein stability and reactivity can be varied within wide limits via incorporation of fluorinated, brominated and unsaturated side chains.					
14. SUBJECT TERMS protein engineering, non-natural amino acids				15. NUMBER OF PAGES 8	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL		

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. Z39-18
298-102

20010608 080

Foreword

This Final Report concerns an ARO-supported program directed toward the development of methods for the incorporation of non-natural amino acids into proteins made in bacterial cells. Our current support for this program was initiated as ARO Grant DAAG-97-1-0386, awarded to the University of Massachusetts (start date September 1, 1997). On September 28, 1998, this grant was re-initiated as DAAG55-98-1-0518 at the California Institute of Technology, following transfer of the PI from UMass to Caltech. The end date on DAAG55-98-1-0518 is September 27, 2000.

Statement of the Problem

Our original proposal outlined three objectives.

1. A systematic exploration of the structures and properties of artificial proteins containing 3-thienylalanine (3-TA), a phenylalanine analogue shown in our prior ARO-supported work to exhibit good translational activity in *Escherichia coli*.
2. Biosynthetic incorporation of *p*-bromophenylalanine (pBF) into artificial proteins, through the use of a bacterial phenylalanyl-tRNA synthetase (PheRS) with relaxed substrate specificity.
3. Development of methods for site-directed insertion of non-natural amino acids (particularly *p*-fluorophenylalanine (pFF)) into proteins *in vivo*.

Summary of Results

Objectives 2 and 3 have been achieved. Furthermore, we have:

1. Demonstrated a new, unanticipated approach to the incorporation of non-natural amino acids into bacterial proteins, via over-expression of the cognate wild-type aminoacyl-tRNA synthetase of the host
2. Determined the rates of activation of a series of methionine analogues by the wild-type *E. coli* methionyl-tRNA synthetase (MetRS)
3. Characterized a mutant form of the *E. coli* MetRS with respect to activation of methionine and its analogues
4. Developed methods for quantitative chemical modification of pBF and applied these methods to crosslinking of engineered artificial proteins
5. Demonstrated stabilization of coiled-coil motifs in leucine-zipper peptides through incorporation of 5,5,5-trifluoroleucine (tfL), a leucine analogue shown in our earlier ARO-supported work to allow engineering of the surface properties of artificial protein films

Each of these developments is described briefly in the following paragraphs.

Biosynthetic incorporation of *p*-bromophenylalanine. *p*-Bromophenylalanine (pBF) does not support protein synthesis in bacterial cells deprived of phenylalanine, owing to the extremely low rate of activation of the brominated amino acid by the wild-type bacterial phenylalanyl-tRNA synthetase (PheRS). In 1991, however, Kast and Hennecke

reported an engineered mutant form (A294G) of this enzyme, in which replacement of an active-site alanine by glycine appeared to allow charging of pBF to tRNA^{Phe}. With ARO support, we have outfitted an *E. coli* strain with PheRS-A294G and demonstrated the use of this strain to achieve efficient incorporation of pBF into a heterologous test protein (mouse dihydrofolate reductase – mDHFR) expressed *in vivo*.

The A294G form of the α -subunit of PheRS was subcloned into the *Pvu*II site of the expression vector pQE15. The resulting plasmid, designated pQE-FS, carries the gene for the target protein DHFR under T5 promoter control. A phenylalanine auxotrophic derivative of *E. coli* strain BL21(DE3), previously constructed in our laboratory with ARO support was used to develop the expression host AF-IQ[pQE-FS]. A control expression host (AF-IQ[pQE15]) did not carry the mutant form of PheRS.

AF-IQ[pQE-FS] and AF-IQ[pQE15] cultures were grown in Phe-supplemented minimal media to OD₆₀₀ of 0.8-1.0, and then shifted to media containing either: i). pBF, ii). Phe (positive control), or iii). no Phe or pBF (negative control). Expression was induced by addition of isopropylthiogalactoside (IPTG) and accumulation of DHFR was assessed by polyacrylamide gel electrophoresis (PAGE). No DHFR was detected in cultures of AF-IQ[pQE15] supplemented with pBF, whereas AF-IQ[pQE-FS] cultures treated similarly afforded protein yields approximately 70% of those found in Phe-supplemented media. Amino acid analysis and NMR spectroscopy showed that as much as 88% of the Phe in the product protein was replaced by pBF under our typical expression conditions.

The development of an efficient means of incorporating pBF into proteins made *in vivo* opens the door to new methods of post-translational modification and structure determination via multiwavelength anomalous diffraction. Furthermore, these results establish the feasibility of engineering organisms through introduction of mutant aminoacyl-tRNA synthetases that can achieve high-level (“uniform”) replacement of natural residues by artificial analogues. Efficient uniform replacement is essential for engineering of the overall physical properties (e.g., solubility, hydrophobic character, etc) of natural and artificial proteins.

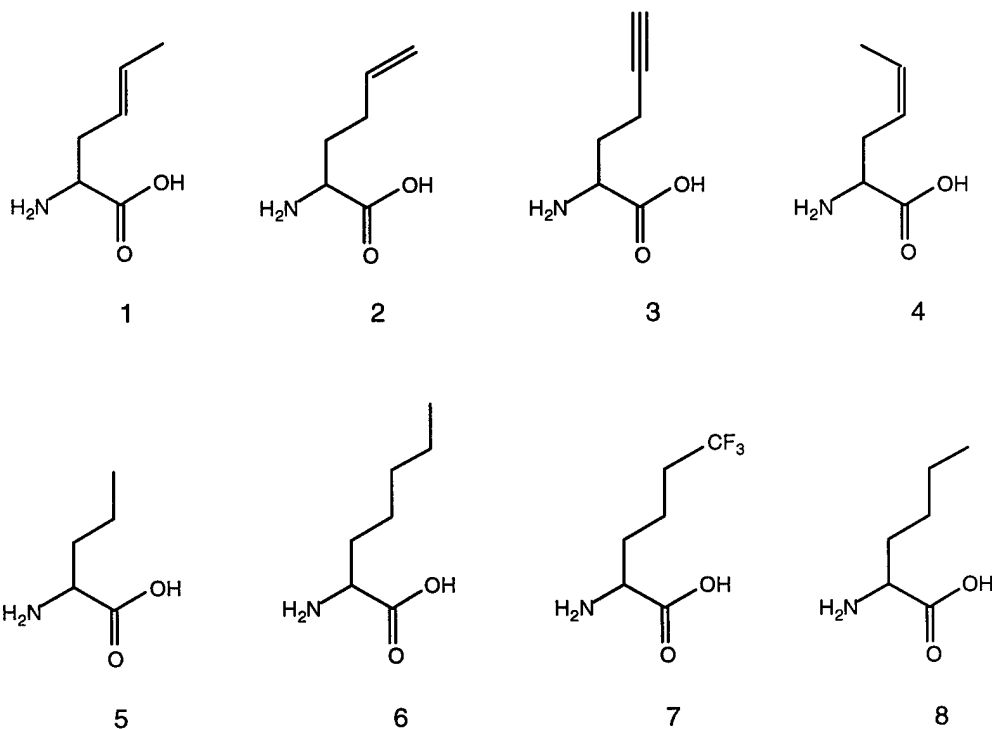
Site-directed insertion of non-canonical amino acids into proteins *in vivo*. Although uniform replacement of natural amino acids by their analogues offers the advantages noted above, there are many problems that require insertion of an analogue at a single site. Site-directed insertion has now been demonstrated for many analogues *in vitro* via chemical misacylation of suppressor tRNAs; however, the scale of such experiments is necessarily small, and an *in vivo* method that accomplishes efficient insertion would bring many advantages. Such a method has been developed and demonstrated during the current funding period.

Our approach involves the introduction of a heterologous (yeast) tRNA/synthetase pair with relaxed amino acid substrate specificity into an *E. coli* expression host. Specifically, a yeast suppressor-tRNA^{Phe}/PheRS pair was expressed in an analogue-resistant *E. coli* strain to direct analogue incorporation at a programmed amber stop codon in the DHFR marker protein. The programmed position could be translated up to 89% as *p*-fluorophenylalanine (pFF), an analogue that is not incorporated into proteins by the host

strain lacking the yeast pair. Depending on expression conditions, pFF incorporation was 10-40-fold higher at the programmed ("stop") position than at a Phe codon engineered into the same site in a related (*Phe5*) allele of DHFR. Protein expression yields were approximately 70% of those obtained for the wild-type marker protein, sufficient to allow routine preparation of protein samples on a scale of hundreds of milligrams.

Over-expression of wild-type MetRS to allow incorporation of non-canonical amino acids. In earlier work directed toward incorporation of unsaturated amino acids into proteins *in vivo*, we found that *trans*-crotylglycine (**1**) would not support measurable levels of protein synthesis in bacterial cultures depleted of methionine. Determination of the rates of activation of a series of methionine analogues, however, indicated that **1** was activated by MetRS roughly half as rapidly as homoallylglycine (**2**), which does support protein synthesis. This result suggested that a modest increase in the MetRS activity of the host might allow efficient incorporation of **1**.

The experiment designed to test this conjecture was similar to that described above for incorporation of pBF, except that the synthetase cloned into the host was the wild-type MetRS. The expression plasmid pQE15 was modified by introduction of a single copy of the MetRS gene, and the resulting plasmid (pQE15-MRS) was used to transform *E. coli* strain B834(DE3). Methionine-depleted media supplemented with **1** supported protein synthesis by cultures bearing the modified plasmid, but not by cultures transformed with pQE15. Protein yields under typical expression conditions were approximately 30% of those obtained in media supplemented with methionine, and N-terminal sequencing showed the level of methionine substitution to be ca. 98%



Activation of methionine analogues by wild-type MetRS. In our initial *in vivo* assays of the translational activities of methionine analogues, we found that **2** and **3** supported protein synthesis while **1** and **4 - 7** did not. (Norleucine (**8**) has been known for many years as a methionine surrogate). We were puzzled, however, by the fact that cultures supplemented with **3** afforded protein yields nearly identical to those obtained with methionine, while **2** gave yields that were reduced by about a factor of 5. The geometry of **3**, with its extended (linear) side-chain terminus, seemed less likely than that of **2** to support rapid activation by MetRS, and we wondered whether protein yields would correlate at all with rates of activation.

They do. Protein yield decreases with decreasing rate of activation of the analogues, although the slower activation of **3** (relative to methionine) is not detrimental to protein yield. These results suggest that activation becomes limiting for protein yield only for those analogues activated more slowly than **3**. It was this result that prompted us to explore over-expression of MetRS as a means of incorporating the very sluggish analogue **1**, as described earlier.

Characterization of a mutant MetRS with respect to activation of methionine analogues. Fourmy and coworkers have reported several mutant forms of *E. coli* MetRS. Of special interest is a comparison of two mutants at position 305 (a tryptophan in the wild-type enzyme). Substitution of alanine to give W305A results in nearly complete inactivation of the enzyme, while the phenylalanine mutant (W305F) shows kinetic behavior almost identical to that of the parent enzyme with respect to activation of methionine. We imagined that the W305F mutant might be characterized by relaxed substrate specificity, owing to a slightly more "open" substrate binding site. We therefore determined the rates of activation of analogues **1 - 8** by MetRS-W305F.

As expected, the mutant is more permissive than the parent; however, the effect is small. For analogue **2** for example, k_{cat}/K_m is four-fold larger for W-305F than for wild-type MetRS. We have not yet made use of this modest increase in the activity of the mutant.

Chemical modification of pBF. One of the objectives of our work on pBF has been the development of new chemistry for post-translational modification of natural and artificial proteins. Aryl bromides undergo a wide variety of metal-catalyzed coupling reactions that are well-suited to the introduction of crosslinks, chromophores, surface-binding sites, etc., to engineered proteins.

We have been especially interested in the Heck, Sonagashira, and Suzuki coupling reactions. By using N-acetyl-pBF (NpBF) as a model substrate, we have identified conditions under which each of these reactions affords quantitative yields of the desired coupling products. All three reactions are readily conducted in aqueous solutions at temperatures mild enough for many protein modification schemes. Although we have not yet conducted a thorough investigation of interference by other amino acid side chains, a preliminary study of the Heck coupling of NpBF with acrylic acid showed interference only by cysteine. We are currently exploring the functional group tolerance of these (and other) coupling reactions of pBF.

Stabilization of coiled-coil peptide dimers through incorporation of tFL. In earlier work supported by ARO, we demonstrated: i). Efficient replacement of leucine by tFL in artificial proteins expressed in bacterial hosts, and ii). marked alteration of the surface properties of protein films containing tFL. More recently, we have extended this work to include examination of the effects of tFL incorporation into leucine-zipper peptides.

Leucine – zipper peptides provide ideal models for the study of secondary, tertiary and quaternary structures in proteins, and are of special relevance to the elucidation of hydrophobic effects on protein stability. Such peptides share a generic heptad repeat, designated – *abcdefg* –, wherein the *d* position is occupied to a large extent by leucine. Hydrophobic interactions drive the formation of coiled-coil dimers, trimers and tetramers in such systems, depending on the details of the peptide sequence.

We have examined the effect of tFL on the stability of the coiled-coil dimer of GCN4-p1, a 33-residue peptide derived from the yeast transcription factor GCN4. We imagined that fluorination of the hydrophobic dimer interface might stabilize the coiled-coil structure with respect to thermal and chemical denaturation.

Both conjectures proved to be correct. The melting temperature of the fluorinated peptide is elevated by 13° C, and the concentration of guanidinium hydrochloride required for 50% unfolding is elevated substantially.

Papers in Peer-Reviewed Journals

R. Furter. “Expansion of the genetic code: Site-directed *p*-fluoro-phenylalanine incorporation in *Escherichia coli*” *Protein Science*, 7, 1 (1998).

N. Sharma, R. Furter, P. Kast and D.A. Tirrell. “Efficient Introduction of Aryl Bromide Functionality into Proteins *in vivo*” *FEBS Lett.*, 467, 37 (2000).

K.L. Kiick and D.A. Tirrell. “Protein Engineering by *In Vivo* Incorporation of Non-natural Amino Acids: Control of Incorporation of Methionine Analogues by Methionyl-tRNA Synthetase” *Tetrahedron*, 56(48), 9487 (2000).

K.L. Kiick, J.C.M. van Hest and D.A. Tirrell. “Expanding the Scope of Protein Biosynthesis by Altering the Methionyl-tRNA Synthetase Activity of a Bacterial Expression Host” *Angew. Chem. Int. Ed.*, 39, 2148 (2000).

Y. Tang, G. Ghirlanda, N. Vaidehi, J. Kua, D.T. Mainz, W.A. Goddard III, W.F. DeGrado and D.A. Tirrell. “Stabilization of Coiled-Coil Peptide Domains by Introduction of Trifluoroleucine” *Biochemistry*, 40, 2790 (2001).

Y. Tang, G. Ghirlanda, W. Petka, T. Nakajima, W.F. DeGrado and D.A. Tirrell. “Fluorinated Coiled-coil Proteins Prepared *in vivo* Display Enhanced Thermal and Chemical Stability.” *Angew. Chem. Int. Ed.*, 40, 1494 (2001).

K.L. Kiick, R. Weberskirch and D.A. Tirrell. "Identification of an Expanded Set of Translationally Active Methionine Analogues in *Escherichia coli*" *FEBS Lett.*, submitted.

Participating Scientific Personnel

Nandita Sharma, Graduate Research Assistant. Ph.D. awarded 2001.

Kristi L. Kiick, Graduate Research Assistant. Ph.D. awarded 2001.

Isaac Carrico, Graduate Research Assistant.

David Flanagan, Graduate Research Assistant.

Yi Tang, Graduate Research Assistant.

Rolf Furter, Visiting Scientist

Report of Inventions

Submitted previously.